

**RECEIVED
CENTRAL FAX CENTER****JAN 12 2005****CASE NO.: 50P4487
Serial No.: 09/932,760
January 12, 2005
Page 2****PATENT
Filed: August 16, 2001****Replacement Paragraphs****Page 1**

Speech recognition systems convert the utterance of a user into digital form and then process the digitized speech in accordance with known algorithms to recognize the words and/or phrases spoken by the user. For example, ~~in co-pending application serial No. (450103-02964)~~ a speech recognition systems [is] have been disclosed wherein the digitized speech is processed to extract sequences of feature sets that describe corresponding speech passages. The speech passage is then recognized by matching the corresponding feature set sequence with the optimal model sequence.

Page 2, paragraph containing line 23:

Such correction and retraining is relatively simple when adapting speech models of a user whose speech matches the data used to train the speech recognition system, because speech from the user and the training group have certain common characteristics. Hence, relatively small modifications to a preset speech model to adapt from those common characteristics are readily achievable, but large deviations are not. Various accents, inflections, pathological speech or other speech features contained in the utterances of such an individual are sufficiently different from the preset speech models as to inhibit successful adaptation retraining of those models. For example, the acoustic subwords pronounced by users whose primary language is not the system target language are quite different from the target language acoustic subwords to which the speech models of typical speech recognition systems are trained. In general, subwords pronounced by "non-native" speakers typically

1168-158.AMD

CASE NO.: 50P4487
Serial No.: 09/932,760
January 12, 2005
Page 3

PATENT
Filed: August 16, 2001

exhibit a transitional hybrid between the primary language of those users and target language subwords. In another example, brain injury, or injury or malformation of the physical speech production mechanism, can significantly impair a speaker's ability to pronounce certain acoustic subwords in conformance with the speaking population at large. A significant subgroup [if] of this speech-impaired population would require the speech models such a system would create.

Page 5, lines 12 and 13:

Figs. 3A-3[B]C constitute a more detailed flow chart depicting the operation of the present invention.

Page 5, first full paragraph of Detailed Description:

Referring now to Fig. 1, there is illustrated a block diagram of the speech recognition system in which the present invention finds ready application. The system includes user sites 10₁, 10₂, ... 10_n, each having a user input 12 and speech recognition apparatus 14; and a remote, central site having a central database 18 and a speech recognition processor module 20 that operates to correct and retrain speech models stored in central database 18. User input 12 preferably includes a suitable microphone or microphone array and supporting hardware of the type that is well known to those of ordinary skill in the art. ~~One example of a suitable user input is described in aforementioned application (attorney docket 450100-02964).~~

Page 9, paragraph containing line 26:

In addition to entering criteria data, the user also enters utterances which are sampled and compared to a predetermined set of stored speech models by speech recognition apparatus 14, as represented by step 34 in Fig. 2. As mentioned previously, the speech recognition

1156-158.AMD

CASE NO.: 50P4487
Serial No.: 09/932,760
January 12, 2005
Page 4

PATENT
Filed: August 16, 2001

apparatus operates in a manner known to those of ordinary skill in the art, ~~such as described in co-pending application (attorney's docket 450100-02964),~~ to sense the user's utterances. The sensed utterance is sampled to extract therefrom identifiable speech features. These speech features are compared to the stored speech models and the optimal match between a sequence of features and the stored models is obtained. It is expected that the best-matched model nevertheless differs from the sampled feature sequence to the extent that an improved set of speech models is needed to optimize recognition performance. As represented by step 36, these models are downloaded from a suitable library, such as a read-only memory device (e.g. a CD-ROM, a magnetic disk, a solid-state memory device, or the like) at the user's site.

Page 11, last full paragraph beginning on line 24:

Referring now to the flow chart shown in Figs. 3A-3[B]C, there is illustrated a more detailed representation of the system operation in accordance with the present invention. Step 62, like step 32 discussed above in conjunction with Fig. 2A, is carried out by the user who enters criteria data by operating user input 12 at, for example, user site 10₁. Thus, as described above, the user enters criteria information, including the primary language spoken by the user, the user's gender, the user's age, the user's height, the user's weight, the age of the user when he first learned the system target language, the number of years the user has spoken the target language and speech samples consisting of registration sentences. Speech recognition apparatus 14 operates by sequentially applying the user's speech samples, one at a time, to a library of stored speech models, as represented by step 64. The operation

1168-158.AMD

CASE NO.: 50P4487
Serial No.: 09/932,760
January 12, 2005
Page 5

PATENT
Filed: August 16, 2001

of the speech recognition apparatus advances to step 66 to find the best match between the given user's speech samples and the stored speech models.

Page 12, last full paragraph beginning on line 22:

In the embodiment depicted in Fig. 3[A]B, the operation of using phonotactics and other conventional speech recognition rules are used in speech recognition apparatus 14 to link, or string acoustic subwords together so as to recognize words, as opposed to this operation being carried out at the central site, described above in conjunction with Figs. 2A-2B. Step 74 depicts this speech recognition operation carried out at the user's site. For example, depending upon the user's class, as determined by his registration of criteria data, the word "six" will be recognized differently from a user whose dialect is from the northern part of United States than from a user whose dialect is from the South. The linking of acoustic subwords from a Northerner may appear as, e.g., "s"--"ih"--"k"--"s"; whereas the linking of acoustic subwords from a Southerner may appear as "s"--"eh"--"k"--"s". Depending upon the registered class of the user, the utterances may or may not be recognized. Then, inquiry 76 (Fig. 3[B]C) is made to determine if the recognized words are correct. For example, if the result of step 74 yields a rejection score within a range of relatively low confidence, the speech recognition apparatus will return to the user, such as by way of a visual or audio cue, the query: "do you mean...?" If the user replies in the negative, thus meaning that the words recognized by step 74 are not correct, inquiry 76 is answered in the negative and the operation of the speech recognition system returns to step 74. Similarly, if the word recognized by step 74 is displayed to the user, either visually or

1168-158.AMD

CASE NO.: 50P4487
Serial No.: 09/932,760
January 12, 2005
Page 6

PATENT
Filed: August 16, 2001

audibly, and the user corrects the displayed word, inquiry 76 is answered in the negative or the user gives up. The system cycles through the loop form of step 74 and inquiry 76 until the inquiry is answered in the affirmative.

1166-138.AMD